Observer Agreement and Accuracy of ^{18}F-Sodium Fluoride PET/CT in the Diagnosis of Bone Metastases in Prostate Cancer

Helle D. Zacho[1,2], Randi F. Fonager[1], Julie B. Nielsen[1], Christian Haarmark[3], Helle W. Hendel[3], Martin B. Johansen[4], Jesper C. Mortensen[5], and Lars J. Petersen[1,2]

[1]Department of Nuclear Medicine, Clinical Cancer Research Center, Aalborg University Hospital, Aalborg, Denmark; [2]Department of Clinical Medicine, Aalborg University, Aalborg, Denmark; [3]Department of Nuclear Medicine, Herlev and Gentofte Hospital, Herlev, Denmark; [4]Unit of Clinical Biostatistics, Aalborg University Hospital, Aalborg, Denmark; and [5]Department of Nuclear Medicine, Regional Hospital West Jutland, Herning, Denmark

Our aim was to evaluate the interobserver agreement in ^{18}F-sodium fluoride (NaF) PET/CT for the detection of bone metastases in patients with prostate cancer (PCa). **Methods:** ^{18}F-NaF PET/CT scans were retrieved from all patients who participated in 4 recent prospective trials. Two experienced observers independently evaluated the ^{18}F-NaF PET/CT scans on a patient level using a 3-category scale (no bone metastases [M0], equivocal for bone metastases, and bone metastases present [M1]) and on a dichotomous scale (M0/M1). In patients with no more than 10 lesions, the location and number of lesions were recorded. On a patient level, the diagnostic performance was calculated using a sensitivity analysis, in which equivocal lesions were handled as M0 as well as M1. **Results:** ^{18}F-NaF PET/CT scans from 219 patients with PCa were included, of whom 129 patients were scanned for primary staging, 67 for biochemical recurrence, and 23 for metastatic castration-resistant PCa. Agreement between the observers was almost perfect on a patient level (3-category unweighted κ = 0.83 ± 0.05, linear weighted κ = 0.90 ± 0.06, and dichotomous κ = 0.91 ± 0.07). On a lesion level (dichotomous scale), the observers agreed on the number and location of bone metastases in 205 (93.6%) patients. In the remaining 14 patients, the readers disagreed on the number of lesions in 13 patients and the location of bone metastases in 1 patient. A final diagnosis of bone metastases was made for 211 of 219 patients. The sensitivity ranged from 0.86 to 0.92, specificity from 0.83 to 0.97, positive predictive value from 0.70 to 0.93, and negative predictive value from 0.94 to 0.96. **Conclusion:** The interobserver agreement on ^{18}F-NaF PET/CT for the detection of bone metastases in patients with PCa was very high among trained observers, both on a patient level and on a lesion level. Moreover, the diagnostic performance of ^{18}F-NaF PET/CT was satisfactory, rendering ^{18}F-NaF PET/CT a robust tool in the diagnostic armamentarium.

Key Words: ^{18}F-sodium-fluoride PET/CT; prostate cancer; interobserver agreement; κ value; bone metastases

Prostate cancer (PCa), the most common cancer in European men (*1*), has a predilection for metastasizing to the bones. A correct diagnosis of bone metastases is fundamental to determine the correct treatment for an individual patient. The European Association of Urology and the National Comprehensive Cancer Network recommend bone scintigraphy for assessment of bone metastases (*2,3*).

With increasing access to PET/CT, several centers have replaced bone scintigraphy with ^{18}F-sodium fluoride (NaF) PET/CT for the assessment of bone metastases. ^{18}F-NaF PET/CT exhibits improved pharmacokinetic properties and a favorable target-to-background ratio compared with methylene diphosphonate, which is frequently used for bone scintigraphy (*4*). Additionally, several studies imply improved diagnostic accuracy by ^{18}F-NaF PET/CT compared with bone scintigraphy (*5–7*).

Despite favorable properties in ^{18}F-NaF PET/CT, imaging evaluation remains a matter of subjectivity and has been labeled the weakest aspect of evaluation (*8,9*). We have previously shown excellent interobserver agreement among trained nuclear medicine physicians for the assessment of bone metastases by planar bone scintigraphy (*10*), but to the best of our knowledge, no such studies have been conducted using ^{18}F-NaF PET/CT.

The aim of the present exploratory study was, primarily, to evaluate the interobserver agreement for ^{18}F-NaF PET/CT findings among experienced observers in various settings of PCa and, second, to assess diagnostic accuracy for bone metastasis in patients with PCa.

MATERIALS AND METHODS

Patients

We included patients with PCa who had participated in 4 prospective studies at our department (*6,11–13*). All patients had undergone ^{18}F-NaF PET/CT as part of the study-related procedures. If more than 1 ^{18}F-NaF PET/CT scan had been conducted, the first was was included in the present analysis. The patients represent a wide range of disease stages, from newly diagnosed PCa and biochemical recurrence after previous curative intent treatment to metastatic castration-resistant PCa.

[18]F-NaF PET/CT

[18]F-NaF PET/CT was conducted in accordance with the guidelines from the Society of Nuclear Medicine and Molecular Imaging (*14*) and the European Association of Nuclear Medicine (*15*) as previously described (*6*). In short, image acquisition was performed in 3-dimensional

mode from the vertex to the mid thigh 30 min after intravenous injection of approximately 200 MBq of ^{18}F-NaF. The images were reconstructed using attenuation correction and an ordered-subset expectation maximization algorithm. Immediately after the PET acquisition, low-dose CT was conducted for use in attenuation correction and anatomic coregistration.

Masked Evaluation of ^{18}F-NaF PET/CT Findings

Each patient's set of ^{18}F-NaF PET/CT images was independently interpreted by 2 board-certified nuclear medicine physicians. Both readers were very experienced (having evaluated more than 2,000 ^{18}F-NaF PET/CT scans at the start of the study). The observers were masked to all clinical information except the PCa diagnosis.

The ^{18}F-NaF PET/CT findings were categorized on a patient level into 1 of 3 categories: no bone metastases (M0), equivocal, and bone metastases (M1). In addition, the observers had to classify the ^{18}F-NaF PET/CT findings dichotomously as bone metastases being present or not present. For patients with 10 or fewer lesions, the observers marked metastatic and equivocal lesions on a schematic drawing of a skeleton to evaluate whether they were considering the same lesions metastatic in each patient. Patients with a metastatic superscan were included in the category with more than 10 bone metastases.

Clinical Data and Final Diagnosis of Metastatic Status

Clinical data were retrieved from the case report forms from the prospective studies. Patients were categorized as having biochemical recurrence and castration-resistant PCa on the basis of the criteria from the European Association of Urology (2,16). In most patients (211/219, 96%), a final diagnosis for the presence or absence of bone metastases was available from the initial study in which the patient participated (6,11–13). In short, in every study a final diagnosis on a patient level was achieved by combining all clinical data, biochemical data, and imaging conducted before inclusion in the study and during follow-up. This process was done by a group of experienced nuclear medicine physicians, a radiologist, and a urologist. In 2 studies, the patients underwent systematic imaging follow-up (6,11). In 1 study, the patients underwent additional MRI (diffusion-weighted MRI), and at least 2 y of follow-up was available for all patients (12). Finally, in 1 study, all patients underwent radical prostatectomy. All patients with postoperative PSA values below 0.1 ng/mL for at least 6 mo after surgery was classified as being without bone metastasis at staging (13).

Statistics

The agreement between the 2 observers was assessed both on the original rating in 1 of the 3 categories and on the dichotomous classification. For the 3-category scale, we used the Cohen κ calculated as unweighted and with a linear weighting, treating the equivocal option as 0.5. The κ was categorized according to the suggestion by Landis and Koch (17). For the dichotomous scale, we calculated κ as well as the positive and negative agreement (p_{pos} and p_{neg}), defined as twice the number of cases for which the 2 observers agreed on a positive or negative rating divided by the sum of the number of cases for which each observer gave a positive or negative rating. Because the p_{pos} and p_{neg} are not simple proportions of independent observations, we calculated percentile-based confidence intervals from a bootstrapping procedure with 1,000 replications for these measures. All other confidence intervals are analytic.

Furthermore, the observer interpretation of the ^{18}F-NaF PET/CT was evaluated against the final diagnosis (used as the reference) by calculating the sensitivity, specificity, positive predictive value, and negative predictive value for each observer and for the consensus rating by the 2 observers. For the 3-category scale, we calculated diagnostic accuracy using sensitivity analysis, with equivocal findings being considered either M0 or M1. Statistical analyses were performed using STATA, version 15 (StataCorp LP).

Approvals

This study complied with the Helsinki II Declaration. All patients provided written informed consent to participate. The studies were approved by the ethical committee (N-20130068, N-20140042, N-20140057, and N-20140080) and by the Danish Data Protection Agency.

RESULTS

Patients

In total, 219 patients were included in the analysis of interobserver agreement. The patients had participated in 4 recent prospective studies (6,11–13), and ^{18}F-NaF PET/CT was performed as a part of the study-related procedures. ^{18}F-NaF-PET/CT was performed at primary staging in 129 patients and at the time of biochemical recurrence after previous treatment with curative intent in 67 patients. Finally, ^{18}F-NaF PET/CT was performed on 23 patients with metastatic castration-resistant PCa (Table 1).

Observer Agreement on a Patient Level

Using the 3-category scale, crude agreement between the 2 observers was seen in 200 of 219 (91%) patients, corresponding to an unweighted κ value of 0.83 ± 0.05 (linear weighted κ was 0.90 ± 0.06). The observers agreed that 131 patients had no bone metastases, that 61 patients had bone metastases, and that the ^{18}F-NaF PET/CT findings were equivocal in 8 patients. The observers disagreed on 19 patients (Supplemental Table 1); in 1 patient, the observers

TABLE 1
Patient Demographics and Clinical Characteristics

Characteristic	Data
Total patients (*n*)	219
Mean age (y)	68 (range, 46–87)
Median PSA (ng/mL)	8.0 (range, 0.2–9,708)
Disease stage at time of ^{18}F-NaF PET/CT (*n*)	
Staging	129
Biochemical recurrence	67
mCRPC	23
Gleason score (*n*)	
<7	17 (8%)
7	121 (55%)
8	31 (14%)
9	38 (17%)
10	2 (1.0%)
Unknown	10 (5%)
T-stage (*n*)	
T1	66 (30%)
T2	86 (39%)
T3	49 (22%)
T4	11 (5%)
Unknown	7 (4%)

PSA = prostate-specific antigen; mCRPC = metastatic castration-resistant PCa.

TABLE 2
Agreement Between 2 Observers on Different Disease Stages of PCa

		Stage of PCa		
Scale	All patients	Staging (*n* = 129)	BCR (*n* = 67)	mCRPC (*n* = 23)
3-category				
Unweighted Cohen κ	0.83 (0.79–0.86)	0.82 (0.74–0.85)	0.70 (0.48–0.78)	1.00
Linear weighted Cohen κ	0.90 (0.87–0.91)	0.89 (0.83–0.97)	0.79 (0.63–0.89)	1.00
Dichotomous				
p_{pos}	0.94 (0.90–0.98)	0.94 (0.87–0.99)	0.86 (0.67–1.00)	1.00
p_{neg}	0.97 (0.95–0.99)	0.97 (0.95–0.99)	0.97 (0.94–1.00)	1.00
Cohen κ	0.91 (0.86–0.97)	0.91 (0.83–0.99)	0.83 (0.65–1.00)	1.00

BCR = biochemical recurrence; mCRPC = metastatic castration-resistant PCa.
Data in parentheses are 95% confidence intervals.

had a 2-category difference (no bone metastases vs. bone metastases), whereas in the remaining 18 patients, the difference in categorization was equivocal versus no bone metastases or equivocal versus bone metastases.

Using the dichotomous scale, the crude agreement increased. The observers agreed on the M0/M1 classification in 211 of 219 (96%) patients (κ = 0.91 ± 0.07). The p_{pos} (M1) was 0.94, and the p_{neg} (M0) was 0.97.

In addition, the agreement was assessed for each disease stage (Table 2). The best observer agreement was found for metastatic castration-resistant PCa (κ = 1.0), whereas the greatest disagreement was observed for biochemical recurrence (unweighted κ = 0.70, linear weighted κ = 0.79). Overall, disagreement was seen in patients with ^{18}F-NaF uptake without corresponding changes on the low-dose CT scan.

Observer Agreement on a Lesion Level (Dichotomous Scale)

In patients with bone metastases according to at least 1 of the observers (*n* = 72), the number of lesions (from 1 to 10) and lesion location were compared between the observers. Thirty-nine patients had more than 10 bone metastases as determined by both observers.

Of the remaining 33 patients, complete agreement on both the number and the location of the lesions was observed in 19 patients (Fig. 1). A difference in the number of metastatic lesions was encountered in 13 patients (Fig. 2; Supplemental Fig. 1; supplemental materials are available at http://jnm.snmjournals.org), of whom 1 patient was categorized as having a metastatic superscan by 1 observer, who also indicated that similar changes may be caused by metabolic bone disease, whereas the other observer indicated benign metabolic bone disease. Finally, the observers indicated different locations of a single lesion of a total of 6 lesions considered to be bone metastases in 1 patient.

### Diagnostic Characteristics of ^{18}F-NaF PET/CT

In most patients (211/219, 96%), a final diagnosis of bone metastases present or absent was established on the basis of clinical and imaging follow-up (6,11–13). For both observers and for the consensus evaluation, the sensitivity, specificity, positive predictive value, and negative predictive value were calculated for the dichotomous evaluation and for the 3-category evaluation, with equivocal results analyzed as negative for bone metastases and then positive for bone

metastases (Table 3). In general, the sensitivity ranged from 0.86 to 0.92, and the specificity ranged from 0.83 to 0.97.

DISCUSSION

To the best of our knowledge, the present study was the first to evaluate interobserver agreement in ^{18}F-NaF PET/CT for the detection of bone metastases. Two experienced masked observers evaluated ^{18}F-NaF PET/CT images for the presence or absence of bone metastases in a large group of patients with PCa and found almost perfect agreement in their assessment of bone metastases. The agreement was investigated across the various stages of PCa with substantial to almost perfect results.

Method evaluation comprised test–retest and observer assessments, and previous studies have primarily focused on test–retest evaluation in ^{18}F-NaF PET/CT, particularly for quantitative parameters such as SUV (18–20), investigated methods for quantification of tumor burden (20), or the inter- and intraobserver variation in the
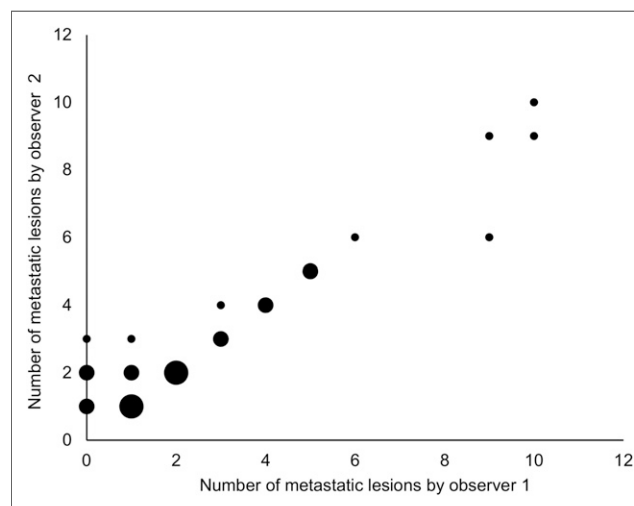
FIGURE 1. Correlation between number of bone metastases by each observer in group of patients with 1–10 bone metastases as determined by at least 1 observer (*n* = 32). Size of dot implies number of identical values it represents. Small dot = 1 patient; medium dot = 2–3 patients; large dot = 4–6 patients.
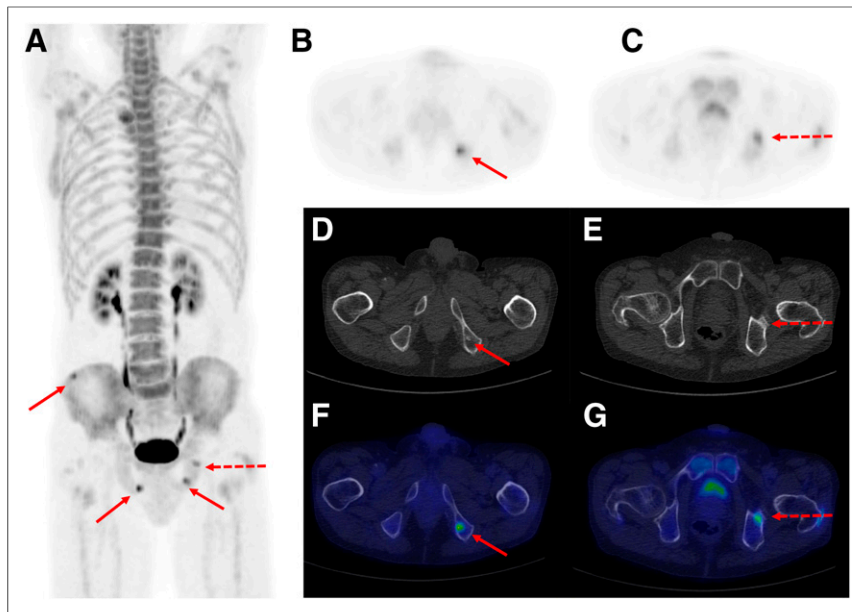
FIGURE 2. (A) Maximum-intensity ^{18}F-NaF PET projection of patient whom observers both deemed to be M1 (correctly, according to reference standard) but disagreed on number of metastases. Observers agreed that patient harbors 3 bone metastases, marked with solid arrows. In addition, 1 observer considered lesion marked by dotted arrow to be metastatic, whereas other observer considered this lesion to be benign. (B, D, and F) One lesion marked by solid arrow on axial ^{18}F-NaF PET image (B) corresponds to sclerotic lesion on CT image (D) and on fused image (F). (C, E, and G) Lesion on which observers disagreed on axial ^{18}F-NaF PET image (C) is shown with corresponding CT image (E) and fused image (G), demonstrating that lesion is near acetabular part of hip.

between the linear weighted κ values, which were 0.72 for bone scintigraphy versus 0.90 for ^{18}F-NaF PET/CT. A similar study using the tracer ^{68}Ga-prostate specific membrane antigen (PSMA) in PET/CT found almost perfect agreement among experienced readers for the presence of bone metastases, with κ values of 0.84 for patients with biochemical recurrence and 0.87 for patients with bone metastases at the time of staging (*22*). Additionally, a large study evaluating PSMA PET/CT in more than 600 patients with biochemical recurrence reported κ values of 0.78 for the assessment of bone metastases, analogous to the present findings (*23*).

The present study was a head-to-head comparison of the number and localization of lesions in patients with up to 10 bone lesions. In most patients, the observers agreed on both the number and the location of bone metastases. In only a minority of patients (*n* = 13) did the number of metastatic lesions differ between the 2 observers, and in only 1 patient did the observers consider different lesions metastatic. This is an important finding in the STAMPEDE era (Systemic Therapy in Advancing or Metastatic Prostate Cancer: Evaluation of Drug Efficacy), in which knowing the number and location of bone metastases is crucial to assigning upfront chemotherapy or abiraterone appropriately (*24,25*) or to offering radiotherapy in men with metastatic PCa who have a low metastatic burden (*26*).

assessment of tumor burden (*21*). Our findings of almost perfect agreement are similar to the previously reported interobserver agreement for planar whole-body bone scintigraphy (*10*), which found a κ value of 0.87 for a dichotomous evaluation of bone scintigraphy. In the study of bone scintigraphy, a 4-category scale for evaluation was available; this might explain the difference observed

Across the range of disease stages, the observer evaluation was compared with a final diagnosis of bone metastases at the patient level. The sensitivity ranged from 0.86 to 0.92, and the specificity

TABLE 3
Diagnostic Accuracy of ^{18}F-NaF PET/CT in 211 Patients with Final Diagnosis for Presence or Absence of Bone Metastases

Scale	M0 by final diagnosis (n = 147)			M1 by final diagnosis (n = 64)			Sensitivity	Specificity	PPV	NPV
	M0	Equivocal	M1	M0	Equivocal	M1				
3-category										
Observer 1	136	7	4	5	4	55	0.86 (0.75–0.93)	0.97 (0.93–0.99)	0.93 (0.84–0.98)	0.94 (0.89–0.97)
							0.92 (0.83–0.97)	*0.93 (0.87–0.96)*	*0.84 (0.74–0.92)*	*0.97 (0.92–0.99)*
Observer 2	122	21	4	5	2	57	0.89 (0.79–0.96)	0.97 (0.93–0.99)	0.93 (0.84–0.98)	0.95 (0.91–0.98)
							0.92 (0.83–0.97)	*0.83 (0.76–0.89)*	*0.70 (0.59–0.80)*	*0.96 (0.91–0.99)*
Consensus	132	11	4	6	2	56	0.88 (0.77–0.94)	0.97 (0.93–0.99)	0.93 (0.84–0.98)	0.95 (0.90–0.98)
							0.91 (0.81–0.96)	*0.90 (0.84–0.94)*	*0.79 (0.68–0.88)*	*0.96 (0.91–0.98)*
Dichotomous										
Observer 1	141		6	7		57	0.89 (0.79–0.96)	0.96 (0.91–0.98)	0.91 (0.80–0.96)	0.95 (0.91–0.98)
Observer 2	137		10	7		57	0.89 (0.79–0.96)	0.93 (0.88–0.97)	0.85 (0.74–0.93)	0.95 (0.91–0.98)
Consensus	142		5	8		56	0.88 (0.77–0.94)	0.97 (0.92–0.99)	0.92 (0.82–0.97)	0.94 (0.90–0.98)

CI = confidence interval; PPV = positive predictive value; NPV = negative predictive value.
Data are for 2 observers (in cases of disagreement, observers reevaluated scans and reached consensus diagnosis) and for optimistic analytic approach (pessimistic approach is in italics). Data in parentheses are 95% confidence intervals.

was 0.83–0.97. The sensitivity reported in the present study was numerically larger than that reported in a recent prospective study by Löfgren et al. (7), whereas the specificity was comparable. Löfgren et al. used a 3-category scale for evaluation, whereas other studies have applied a dichotomous scale (27,28) with results comparable to our dichotomous evaluation. A review of ^{18}F-NaF PET/CT by Wondergem et al. showed a sensitivity of 89% and specificity of 91% for the detection of bone metastases on a patient level (29); these numbers are in line with our data. However, a few studies have reported moderate specificity for ^{18}F-NaF PET/CT due to uptake in benign degenerative and inflammatory lesions. Furthermore, the present diagnostic accuracy for ^{18}F-NaF PET/CT was comparable to similar studies using PSMA PET/CT for bone metastases (23,30).

The use of a 3-category scale provides for the equivocal option and may thus resemble everyday clinical practice more than the dichotomous option does. We found that the proportion of equivocal findings by consensus was considerably lower in the present study of ^{18}F-NaF PET/CT than that observed in a similar study using planar bone scintigraphy, in which approximately 1 in 4 patients had equivocal findings (31). The proportion of equivocal findings on ^{18}F-NaF PET/CT resembles previously published findings in studies that added SPECT/CT to planar bone scintigraphy (6,32,33), as well as the reported proportion of equivocal findings in previous diagnostic test accuracy studies on ^{18}F-NaF PET/CT (5,7). In contrast, the proportion of equivocal ^{18}F-NaF PET/CT findings in the present study was much lower than that stated by the U.S. National PET Registry study, which reported equivocal ^{18}F-NaF PET/CT results in 15% of the patients (34).

To ensure the independence and masking of the observers, they were not involved in recruiting or managing patients in any way. The observers had extensive experience with ^{18}F-NaF PET/CT from having served at a high-volume center for many years. The high level of agreement might in part be due to their vast experience, and the experience level may render the results less generalizable to readers with less experience and to centers with a low volume of ^{18}F-NaF PET/CT-scans. To the best of our knowledge, no studies have investigated the impact of reader experience on agreement or accuracy with ^{18}F-NaF PET/CT. The number of ^{18}F-NaF PET/CT scans one would have to read to reach this expert assessment level remains unknown; in PSMA PET/CT this number is considered to be around 300 (22). Whether the number could be lower for ^{18}F-NaF PET/CT if the observer has extensive experience in bone scintigraphy with SPECT/CT remains undetermined.

According to the U.S. National PET Registry study, ^{18}F-NaF PET/CT changed management in approximately 40% of the patients, and this percentage was reduced to 12%–16% in patients for whom other imaging modalities were planned (34). In a study by Gauthé et al., ^{18}F-NaF PET/CT changed management in 7% of patients at initial staging (35). However, the improved diagnostic properties of ^{18}F-NaF PET/CT or the ^{18}F-NaF PET/CT-induced change in patient management has not yet shown improvement in patient-related outcomes. We recently showed that in patients with negative bone scintigraphy results who underwent radical prostatectomy, ^{18}F-NaF PET/CT did not add prognostic value in terms of the outcome after radical prostatectomy (13). There is a lack of studies showing that ^{18}F-NaF PET/CT improves patient-related outcome, which might be one of the reasons why ^{18}F-NaF PET/CT is not recommended by the international guidelines for assessment of bone metastases.

CONCLUSION

The present study demonstrated almost perfect agreement between 2 observers in using ^{18}F-NaF PET/CT images to evaluate bone metastases in PCa. Likewise, the ^{18}F-NaF PET/CT agreement on a lesion level was substantial, and the diagnostic accuracy was satisfactory, rendering ^{18}F-NaF PET/CT a robust tool for the assessment of bone metastases in PCa. Future studies on ^{18}F-NaF PET/CT should focus on the patient-related outcome to evaluate whether the advantageous properties of ^{18}F-NaF PET/CT are reflected in patient-related outcomes.

DISCLOSURE

ACKNOWLEDGMENT

KEY POINTS

QUESTION: What is the level of interobserver agreement on ^{18}F-NaF PET/CT for the detection of bone metastases in patients with PCa?

PERTINENT FINDINGS: Two experienced observers independently evaluated ^{18}F-NaF PET/CT scans from 4 prospective trials for the detection of bone metastases in 219 patients with PCa across primary staging, biochemical recurrence, and metastatic castration-resistant PCa. Excellent agreement was seen both on a patient level and on a lesion level. Furthermore, satisfactory diagnostic accuracy was seen when the findings were compared with a final diagnosis.

IMPLICATIONS FOR PATIENT CARE: ^{18}F-NaF PET/CT has excellent interobserver agreement and is a robust tool for the detection of bone metastases in patients with PCa.

REFERENCES

1. Taitt HE. Global trends and prostate cancer: a review of incidence, detection, and mortality as influenced by race, ethnicity, and geographic location. *Am J Mens Health.* 2018;12:1807–1823.
2. Mottet N, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent. *Eur Urol.* 2017;71:618–629.
3. Mohler JL, Antonarakis ES, Armstrong AJ, et al. Prostate cancer, version 2.2019, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw.* 2019; 17:479–505.
4. Czernin J, Satyamurthy N, Schiepers C. Molecular mechanisms of bone ^{18}F-NaF deposition. *J Nucl Med.* 2010;51:1826–1829.
5. Even-Sapir E, Metser U, Mishani E, Lievshitz G, Lerman H, Leibovitch I. The detection of bone metastases in patients with high-risk prostate cancer: ^{99m}Tc-MDP planar bone scintigraphy, single- and multi-field-of-view SPECT, ^{18}F-fluoride PET, and ^{18}F-fluoride PET/CT. *J Nucl Med.* 2006;47:287–297.
6. Fonager RF, Zacho HD, Langkilde NC, et al. Diagnostic test accuracy study of ^{18}F-sodium fluoride PET/CT, ^{99m}Tc-labelled diphosphonate SPECT/CT, and planar bone scintigraphy for diagnosis of bone metastases in newly diagnosed, high-risk prostate cancer. *Am J Nucl Med Mol Imaging.* 2017;7:218–227.

7. Löfgren J, Mortensen J, Rasmussen SH, et al. A prospective study comparing 99mTc-hydroxyethylene-diphosphonate planar bone scintigraphy and whole-body SPECT/CT with 18F-fluoride PET/CT and 18F-fluoride PET/MRI for diagnosing bone metastases. *J Nucl Med.* 2017;58:1778–1785.

8. Brealey S, Westwood M. Are you reading what we are reading? The effect of who interprets medical images on estimates of diagnostic test accuracy in systematic reviews. *Br J Radiol.* 2007;80:674–677.

9. Robinson PJ. Radiology's Achilles' heel: error and variation in the interpretation of the Röntgen image. *Br J Radiol.* 1997;70:1085–1098.

10. Zacho HD, Manresa JA, Mortensen JC, Bertelsen H, Petersen LJ. Observer agreement and accuracy in the evaluation of bone scans in newly diagnosed prostate cancer. *Nucl Med Commun.* 2015;36:445–451.

11. Fonager RF, Zacho HD, Langkilde NC, et al. Prospective comparative study of 18F-sodium fluoride PET/CT and planar bone scintigraphy for treatment response assessment of bone metastases in patients with prostate cancer. *Acta Oncol.* 2018;57:1063–1069.

12. Zacho HD, Nielsen JB, Afshar-Oromieh A, et al. Prospective comparison of 68Ga-PSMA PET/CT, 18F-sodium fluoride PET/CT and diffusion weighted-MRI at for the detection of bone metastases in biochemically recurrent prostate cancer. *Eur J Nucl Med Mol Imaging.* 2018;45:1884–1897.

13. Zacho H, Jochumsen MR, Langkilde NC, et al. No added value of 18F-sodium fluoride PET/CT for the detection of bone metastases in patients with newly diagnosed prostate cancer with normal bone scintigraphy. *J Nucl Med.* May 30, 2019 [Epub ahead of print].

14. Segall G, Delbeke D, Stabin MG, et al. SNM practice guideline for sodium 18F-fluoride PET/CT bone scans 1.0. *J Nucl Med.* 2010;51:1813–1820.

15. Beheshti M, Mottaghy FM, Paycha F, et al. 18F-NaF PET/CT: EANM procedure guidelines for bone imaging. *Eur J Nucl Med Mol Imaging.* 2015;42:1767–1777.

16. Cornford P, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG guidelines on prostate cancer. Part II: treatment of relapsing, metastatic, and castration-resistant prostate cancer. *Eur Urol.* 2017;71:630–642.

17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.

18. Kurdziel KA, Shih JH, Apolo AB, et al. The kinetics and reproducibility of 18F-sodium fluoride for oncology using current PET camera technology. *J Nucl Med.* 2012;53:1175–1184.

19. Lin C, Bradshaw T, Perk T, et al. Repeatability of quantitative 18F-NaF PET: a multicenter study. *J Nucl Med.* 2016;57:1872–1879.

20. Rohren EM, Etchebehere EC, Araujo JC, et al. Determination of skeletal tumor burden on 18F-fluoride PET/CT. *J Nucl Med.* 2015;56:1507–1512.

21. Lapa P, Marques M, Costa G, Iagaru A, Pedroso de Lima J. Assessment of skeletal tumour burden on 18F-NaF PET/CT using a new quantitative method. *Nucl Med Commun.* 2017;38:325–332.

22. Fendler WP, Calais J, Allen-Auerbach M, et al. 68Ga-PSMA-11 PET/CT interobserver agreement for prostate cancer assessments: an international multicenter prospective study. *J Nucl Med.* 2017;58:1617–1623.

23. Fendler WP, Calais J, Eiber M, et al. Assessment of 68Ga-PSMA-11 PET accuracy in localizing recurrent prostate cancer: a prospective single-arm clinical trial. *JAMA Oncol.* 2019;5:856–863.

24. James ND, Sydes MR, Clarke NW, et al. Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results from an adaptive, multiarm, multistage, platform randomised controlled trial. *Lancet.* 2016;387:1163–1177.

25. James ND, de Bono JS, Spears MR, et al. Abiraterone for prostate cancer not previously treated with hormone therapy. *N Engl J Med.* 2017;377:338–351.

26. Parker CC, James ND, Brawley CD, et al. Radiotherapy to the primary tumour for newly diagnosed, metastatic prostate cancer (STAMPEDE): a randomised controlled phase 3 trial. *Lancet.* 2018;392:2353–2366.

27. Langsteger W, Balogova S, Huchet V, et al. Fluorocholine (18F) and sodium fluoride (18F) PET/CT in the detection of prostate cancer: prospective comparison of diagnostic performance determined by masked reading. *Q J Nucl Med Mol Imaging.* 2011;55:448–457.

28. Beheshti M, Vali R, Waldenberger P, et al. Detection of bone metastases in patients with prostate cancer by 18F fluorocholine and 18F fluoride PET-CT: a comparative study. *Eur J Nucl Med Mol Imaging.* 2008;35:1766–1774.

29. Wondergem M, van der Zant FM, van der Ploeg T, Knol RJ. A literature review of 18F-fluoride PET/CT and 18F-choline or 11C-choline PET/CT for detection of bone metastases in patients with prostate cancer. *Nucl Med Commun.* 2013;34:935–945.

30. Pyka T, Okamoto S, Dahlbender M, et al. Comparison of bone scintigraphy and 68Ga-PSMA PET for skeletal staging in prostate cancer. *Eur J Nucl Med Mol Imaging.* 2016;43:2114–2121.

31. Zacho HD, Barsi T, Mortensen JC, et al. Prospective multicenter study of bone scintigraphy in consecutive patients with newly diagnosed prostate cancer. *Clin Nucl Med.* 2014;39:26–31.

32. Palmedo H, Marx C, Ebert A, et al. Whole-body SPECT/CT for bone scintigraphy: diagnostic value and effect on patient management in oncological patients. *Eur J Nucl Med Mol Imaging.* 2014;41:59–67.

33. Zacho HD, Manresa JAB, Aleksyniene R, et al. Three-minute SPECT/CT is sufficient for the assessment of bone metastasis as add-on to planar bone scintigraphy: prospective head-to-head comparison to 11-min SPECT/CT. *EJNMMI Res.* 2017;7:1.

34. Hillner BE, Siegel BA, Hanna L, Duan F, Shields AF, Coleman RE. Impact of 18F-fluoride PET in patients with known prostate cancer: initial results from the National Oncologic PET Registry. *J Nucl Med.* 2014;55:574–581.

35. Gauthé M, Aveline C, Lecouvet F, et al. Impact of sodium 18F-fluoride PET/CT, 18F-fluorocholine PET/CT and whole-body diffusion-weighted MRI on the management of patients with prostate cancer suspicious for metastasis: a prospective multicentre study. *World J Urol.* 2019;37:1587–1595.